ELSEVIER

Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT®

Journal of Computational and Applied Mathematics 161 (2003) 41–65

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

Towards the resolution of the Gibbs phenomena

Bernie D. Shizgal*,[1], Jae-Hun Jung

Institute of Applied Mathematics, University of British Columbia, 6356 Agricultural Road, Vancouver, BC, Canada V6T 1Z1

Received 4 February 2003; received in revised form 14 April 2003

Abstract

It is well known that the expansion of an analytic nonperiodic function on a finite interval in a Fourier series leads to spurious oscillations at the interval boundaries. This result is known as the Gibbs phenomenon. The present paper introduces a new method for the resolution of the Gibbs phenomenon which follows on the reconstruction method of Gottlieb and coworkers (SIAM Rev. 39 (1997) 644) based on Gegenbauer polynomials orthogonal with respect to weight function $(1-x^2)^{\lambda-1/2}$. We refer to their approach as the direct method and to the new methodology as the inverse method. Both methods use the finite set of Fourier coefficients of some given function as input data in the re-expansion of the function in Gegenbauer polynomials or in other orthogonal basis sets. The finite partial sum of the new expansion provides a spectrally accurate approximation to the function. In the direct method, this requires that certain conditions are met concerning the parameter λ in the weight function, the number of Fourier coefficients, N and the number of Gegenbauer polynomials, m. We show that the new inverse method can give exact results for polynomials independent of λ and with $m=N$. The paper presents several numerical examples applied to a single domain or to subdomains of the main domain so as to illustrate the superiority of the inverse method in comparison with the direct method.

Keywords: Gibbs phenomenon; Spectral methods

1. Introduction

Spectral and pseudo-spectral methods [3,9,19,23] are based on the expansion of a function in a basis set of functions orthogonal with respect to some weight function on a specified interval.

* Corresponding author. Tel.: +6048223997; fax: 6048222847.
E-mail addresses: shizgal@chem.ubc.ca (B.D. Shizgal), jung@math.ubc.ca (J.-H. Jung).
[1] Also with the Department of Chemistry, University of British Columbia, Main Mall 2036, Vancouver, BC, Canada V6T 1Z1.

0377-0427/03/$ - see front matter
doi:10.1016/S0377-0427(03)00500-4

The function may be the solution of some partial differential equation or could be some particular function of interest. The basis sets most commonly used in the application of spectral methods are the Fourier sine and cosine functions for periodic problems, and Chebyshev or Legendre polynomials for nonperiodic problems. For certain problems, nonclassical basis functions have proven very useful [5,20], in particular for the infinite and semi-infinite domains. It has been recognized for some time that spectral methods can provide a very high order solution to a differential equation provided the solution is smooth and does not exhibit shock-like behaviour. Consequently, it has long been considered that spectral methods may not be useful in the solution of hyperbolic equations that arise in compressible fluid flow. However, considerable progress has been made in overcoming these difficulties [12], and the results of this paper may also provide significant advances in this direction.

The basic problem in the application of spectral methods to hyperbolic systems is that the solutions can change rapidly in the vicinity of a shock. The expansion of such nonsmooth, nonperiodic functions in a Fourier series yields the classic Gibbs phenomena [15,17]. The function, represented by a finite Fourier series, exhibits spurious oscillations near the region where the function is not smooth or discontinuous. These oscillations do not diminish as the number of terms in the expansion is increased. It is a classic textbook exercise to show that with an increase in the number of terms in the Fourier series the oscillations approach a discontinuity but do not diminish in amplitude [17]. The resolution of this difficulty also has important applications to reducing the effects of noise in image reconstruction [2].

The efficient use of spectral methods for problems of this type requires a procedure to filter the high frequency components that give rise to the oscillations. There have been several recent papers proposing different techniques to overcome the spurious oscillations associated with the Gibbs phenomenon and the restoration of spectral accuracy. Tadmor and Tanner [22] have developed adaptive mollifiers for the high resolution recovery of smooth data from its spectral information. Driscoll and Fornberg [8] have discussed the use of a Padé based algorithm for overcoming the Gibbs phenomena. A Padé method was also employed by March and Barone [18]. Adomaitis [1] has presented a spectral filtering to reduce the Gibbs oscillations. Over the past decade, Gottlieb and co-workers [14] have developed a method based on the re-expansion of the finite Fourier representation of a function in Gegenbauer polynomials, $C_\ell^\lambda(x)$. The objective of the present paper is to introduce an alternate methodology, referred to as the inverse method, that is closely related to the Gegenbauer reconstruction. This terminology is made clear in the presentation in Section 3, where the inverse method is described.

The Gegenbauer polynomials are a subset of the Jacobi polynomials and satisfy the orthogonality relation

$$\frac{1}{h_\ell^\lambda}\int_{-1}^{1}(1-x^2)^{\lambda-1/2}C_\ell^\lambda(x)C_k^\lambda(x)\,\mathrm{d}x=\delta_{\ell,k}. \tag{1}$$

The normalization h_ℓ^λ is defined later. The parameter λ in the weight function plays a crucial role in the Gegenbauer reconstruction procedure. Gottlieb and co-workers have provided a detailed numerical analysis of this procedure and have shown that with some appropriate choice of the parameter λ, that is related to the number of terms in the original Fourier series, the expansion of the function converges exponentially, that is, spectral accuracy is recovered. The details of this approach is given in Section 2. We refer to their approach as the direct method.

We compare the two methods in Section 5 by their application to several numerical examples. We also include the use of other orthogonal polynomial sets instead of the Gegenbauer polynomials. In Section 6, we consider the application of the inverse method to the resolution of the Gibbs phenomenon on subintervals of the main domain.

2. Gegenbauer reconstruction procedure: the direct method

We consider a function, $f(x)$, on the interval $[-1,1]$ expanded in a finite Fourier series

$$f_N(x) = a_0 + \sum_{k=1}^{N} [a_k \cos(k\pi x) + b_k \sin(k\pi x)] \tag{2}$$

where the Fourier coefficients are given by

$$a_0 = \frac{1}{2} \int_{-1}^{1} f(x)\,\mathrm{d}x,$$

$$a_k = \int_{-1}^{1} f(x)\cos(k\pi x)\,\mathrm{d}x,$$

$$b_k = \int_{-1}^{1} f(x)\sin(k\pi x)\,\mathrm{d}x. \tag{3}$$

The approximation $f_N(x)$ to the function $f(x)$ will exhibit the Gibbs phenomenon at the interval boundaries or near interior points where the function is not smooth. There are many examples in standard textbooks [13,17]. We consider limited input data consisting of the first N *exact* Fourier coefficients a_k and b_k. The basic objective of the reconstruction procedure is to recover the function $f(x)$ from the input data. We also note the expansion of $f(x)$ in the Gegenbauer polynomials, that is,

$$f(x) = \sum_{\ell=0}^{\infty} g_\ell C_\ell^\lambda(x), \tag{4}$$

where the g_ℓ coefficients depend on λ and are the exact Gegenbauer coefficients given by

$$g_\ell = \frac{1}{h_\ell^\lambda} \int_{-1}^{1} (1-x^2)^{\lambda-1/2} C_\ell^\lambda(x) f(x)\,\mathrm{d}x \tag{5}$$

and

$$h_\ell^\lambda = \frac{2^{1-2\lambda}\pi\Gamma(\ell+2\lambda)}{\ell!(\ell+\lambda)[\Gamma(\lambda)]^2}.$$

The Gegenbauer reconstruction procedure involves the re-expansion of $f_N(x)$, Eq. (2), in a finite series of Gegenbauer polynomials, that is,

$$\hat{f}_m(x) = \sum_{\ell=0}^{m} \hat{g}_\ell C_\ell^\lambda(x), \tag{6}$$

where the (approximate) Gegenbauer coefficients are given by

$$\hat{g}_\ell = \frac{1}{h_\ell^\lambda} \int_{-1}^{1} (1-x^2)^{\lambda-1/2} C_\ell^\lambda(x) f_N(x)\,\mathrm{d}x. \tag{7}$$

It is important to notice that the summation in Eq. (6) is up to $\ell = m$ and in Eq. (2) it is up to $k = N$, hence we denote the reconstructed function by $\hat{f}_m(x)$. An important aspect of the reconstruction procedure is that the summation in Eq. (6) is truncated at some sufficiently *small* value of $m < N$. If this expansion is extended indefinitely, the result would be an increasingly more accurate representation of $f_N(x)$ and one would recover the Gibbs phenomenon. An essential ingredient of the reconstruction procedure is that m is not too large. This will be made more precise in later discussions.

The parameter λ that defines the Gegenbauer polynomials plays a very important role. In the detailed numerical analysis described in [14,15], it was demonstrated that a sufficient condition for spectral convergence of Eq. (6) versus N is that $\lambda = m = \beta N$, where $\beta = 2\pi/27 = 0.2327$. In most of the benchmarks of the direct method that have been described, the choice $\lambda = m = N/4$ is made. We do not repeat any of the detailed analysis presented in the previous papers. In this paper, we consider λ as an adjustable parameter and study its role in the Gegenbauer reconstruction procedure. We also consider different orthogonal polynomial basis functions with parameters analogous to λ in the weight function.

The reconstruction procedure is a basis set transformation, that is the Fourier sine and cosine basis functions are expanded in the Gegenbauer polynomials. We therefore consider the expansions,

$$\begin{aligned} \sin(k\pi x) &= \sum_{\ell=1}^{\infty} S_{k,\ell} C_\ell^\lambda(x), \\ \cos(k\pi x) &= \sum_{\ell=0}^{\infty} W_{k,\ell} C_\ell^\lambda(x), \end{aligned} \tag{8}$$

where the expansion coefficients are as follows:

$$\begin{aligned} S_{k,\ell} &= \frac{1}{h_\ell^\lambda} \int_{-1}^{1} (1-x^2)^{\lambda-1/2} C_\ell^\lambda(x) \sin(k\pi x)\,\mathrm{d}x, \\ W_{k,\ell} &= \frac{1}{h_\ell^\lambda} \int_{-1}^{1} (1-x^2)^{\lambda-1/2} C_\ell^\lambda(x) \cos(k\pi x)\,\mathrm{d}x, \end{aligned} \tag{9}$$

which can be evaluated analytically as follows:

$$\begin{aligned} S_{k,\ell} &= (-1)^{(\ell+1)/2+1} \left(\frac{2}{\pi k}\right)^\lambda (\ell+\lambda)\Gamma(\lambda) J_{\ell+\lambda}(\pi k), \\ W_{k,\ell} &= (-1)^{\ell/2} \left(\frac{2}{\pi k}\right)^\lambda (\ell+\lambda)\Gamma(\lambda) J_{\ell+\lambda}(\pi k), \end{aligned} \tag{10}$$

where $J_\nu(x)$ is the Bessel function. With Eq. (8) in Eq. (2), we have that

$$\hat{g}_\ell = a_0 + \sum_{k=1}^{N} [a_k W_{k,\ell} + b_k S_{k,\ell}]. \tag{11}$$

The essential aspect of this basis set transformation is that the projection of the higher order Fourier basis functions onto the Gegenbauer polynomials is small. These projections are precisely the coefficients $S_{k,\ell}$ and $W_{k,\ell}$ defined by Eq. (9). The work by Gottlieb and co-workers was based on a detailed numerical analysis of these coefficients and their relation to the *Gibbs condition* as defined and discussed elsewhere [12,14].

In the interest of simplicity and clarity, we restrict the ensuing discussion to odd functions on the interval $[-1,1]$. The generalization to functions of arbitrary symmetry is straightforward. For odd functions, $f(-x) = -f(x)$, Eq. (11) reduces to,

$$\hat{g}_\ell = \sum_{k=1}^{N} b_k S_{k,\ell}. \tag{12}$$

The approximate Gegenbauer coefficients are then determined with Eq. (12), and the reconstructed function with Eq. (6).

3. Alternate reconstruction method: the inverse method

We consider a modification of the Gottlieb reconstruction procedure by noting that the desired representation of the function is a *finite* expansion of Gegenbauer (or other) polynomials, that is

$$\tilde{f}_m(x) = \sum_{\text{odd}\,\ell}^{m} \tilde{g}_\ell C_\ell^\lambda(x). \tag{13}$$

For simplicity of presentation, we again restrict the discussion to odd functions. The generalization to other functions is straightforward. In this new approach, instead of expanding the Fourier sines and cosines in the Gegenbauer basis as in Eq. (8), the approximate $\tilde{g}_\ell$ coefficients in Eq. (13) are determined by considering the representation of $\tilde{f}_m(x)$ in a Fourier sine series and then projecting out each Fourier mode. The main idea in this approach is that the sought after representation of the function, Eq. (13), is a finite polynomial. Therefore, we consider the correspondence

$$\sum_{\text{odd}\,\ell}^{\infty} \tilde{g}_\ell C_\ell^\lambda(x) = \sum_{k=1}^{N} b_k \sin(k\pi x). \tag{14}$$

From the orthogonality of the Fourier sine basis functions, the first m $\tilde{g}_\ell$ coefficients are given by

$$\sum_{\text{odd}\,\ell}^{m} \tilde{g}_\ell T_{\ell,k} = b_k, \tag{15}$$

where

$$T_{\ell,k} = \int_{-1}^{1} C_\ell^\lambda(x) \sin(k\pi x)\, \mathrm{d}x. \tag{16}$$

Eq. (15) defines the approximate $\tilde{g}_\ell$ coefficients that appear in Eq. (13). The matrix $T_{\ell,k}$ is the transformation from the Gegenbauer space $\{g_\ell\}$ to the Fourier space $\{b_k\}$, whereas $S_{k,l}$ is the transformation from the Fourier space to the Gegenbauer space. The approximate Gegenbauer coefficients, $\tilde{g}_\ell$, are given by the inversion of Eq. (15) with $m = N$, and implicitly depend on m and λ. We refer to this approach as the inverse method and the original Gegenbauer reconstruction introduced by Gottlieb and co-workers as the direct method. We show later that if $f(x)$ is a polynomial of degree m, the inversion of Eq. (15) gives m exact $\tilde{g}_\ell$ coefficients from m exact b_k Fourier coefficients, whereas the direct method based on Eq. (12) does not yield exact results in the same manner.

The projection of Eq. (14) onto the Fourier basis gives Eq. (15) whereas the projection onto the Gegenbauer basis gives Eq. (12) with $\tilde{g}_\ell$ replaced with $\hat{g}_\ell$. The relationship between the approximate Gegenbauer coefficients in the two methods is obtained by substituting Eq. (15) into Eq. (12), that is

$$\hat{g}_\ell = \sum_{\text{odd } \ell'}^{m} U_{\ell,\ell'} \tilde{g}_{\ell'}, \tag{17}$$

where the matrix $\mathbf{U} = \mathbf{T} \cdot \mathbf{S}$ is a unit matrix only in the limit $N \to \infty$, that is

$$\begin{aligned} U_{\ell,\ell'}^{(s)} &= \sum_{k=1}^{N} T_{\ell,k} S_{k,\ell'} \\ &= \frac{1}{h_{\ell'}^\lambda} \int_{-1}^{1} \int_{-1}^{1} (1-x^2)^{\lambda-1/2} C_{\ell'}^\lambda(x) \left[\sum_{k=1}^{N} \sin(k\pi x) \sin(k\pi y) \right] C_\ell^\lambda(y)\, \mathrm{d}x\, \mathrm{d}y \\ &\to \delta_{\ell,\ell'} \quad \text{as } N \to \infty. \end{aligned} \tag{18}$$

The transformation from the Fourier basis to the polynomial basis is not equivalent to the transformation from the polynomial basis to the Fourier basis, unless an infinite number of Fourier basis functions are retained. The equivalence of the two transformations, for $N \to \infty$, arises from the completeness relation [16],

$$\lim_{N\to\infty} \sum_{k=1}^{N} \sin(k\pi x) \sin(k\pi y) = \delta(x-y) \tag{19}$$

and the orthogonality of the $C_\ell^\lambda(x)$, Eq. (1). The representation of the delta function in Eq. (19) is understood in the context of an integration, that is, $\int f(y)\delta(x-y)\, \mathrm{d}y = f(x)$. It can be shown that the sum over k in Eq. (19) leads to the familiar Dirichlet kernel which is a representation of the delta function [4] in the limit $N \to \infty$. We also define the corresponding matrix $U_{\ell,\ell'}^{(c)}$ that arises

from the Fourier cosine series, that is,

$$U^{(c)}_{\ell,\ell'} = \sum_{k=0}^{N} V_{\ell,k} W_{k,\ell'}$$

$$= \frac{1}{h^{\lambda}_{\ell'}} \int_{-1}^{1}\int_{-1}^{1} (1-x^2)^{\lambda-1/2} C^{\lambda}_{\ell'}(x) \left[\sum_{k=0}^{N} \frac{1}{c_k} \cos(k\pi x)\cos(k\pi y)\right] C^{\lambda}_{\ell}(y)\,\mathrm{d}x\,\mathrm{d}y$$

$$\to \delta_{\ell,\ell'} \quad \text{as } N \to \infty, \tag{20}$$

where

$$V_{\ell,k} = \int_{-1}^{1} C^{\lambda}_{\ell}(x)\cos(k\pi x)\,\mathrm{d}x \tag{21}$$

and $c_k = 2$ for $k = 0$ and $c_k = 1$ for $k \neq 0$.

The *truncation error*, $\mathrm{TE}(\lambda, m, N)$, defined in [15] can be written in terms of $\mathbf{U}$, that is,

$$\mathrm{TE}(\lambda, m, N) = \max_{-1\leqslant x\leqslant 1} \left| \sum_{l'=0}^{m} (g_{l'} - \hat{g}_{l'}) C^{\lambda}_{l'} \right|$$

$$= \max_{-1\leqslant x\leqslant 1} \left| \sum_{l'=0}^{m} \left(g_{l'} - \sum_{l=0}^{\infty} g_l U_{ll'} \right) C^{\lambda}_{l'} \right|.$$

Thus, we find that the truncation error of the direct method is mainly determined by the nature of $U_{ll'}$. If $U_{ll'} = \delta_{ll'}$ we obtain an *exact* reconstruction. However, as we have shown $\mathbf{U}$ is not a unit matrix. Now we observe that the source of the truncation error of the direct method lies in $\mathbf{U}$. From the definition of $\mathbf{U}$ we know that this is because one projects $f(x)$ onto the Fourier space and again back onto the Gegenbauer space. These projections do not commute, i.e., $S \cdot T \neq T \cdot S$ and one does not obtain the completeness relation, Eq. (19) for the finite N. We apply the direct method, Eq. (12), and the inverse method, Eq. (15), to the resolution of the Gibbs phenomenon to several examples in Section 5. In previous studies of the direct method, these test functions have often been low order polynomials. We demonstrate in the next section that the inverse method is exact for polynomials.

4. The inverse method is exact for polynomials

It is useful to note the explicit expressions for the lower-order Gegenbauer polynomials. These are:

$$C^{\lambda}_0(x) = 1,$$

$$C^{\lambda}_1(x) = 2\lambda x,$$

$$C^{\lambda}_2(x) = 2\lambda(\lambda+1)x^2 - \lambda,$$

$$C^{\lambda}_3(x) = \frac{4\lambda}{3}(\lambda+1)(\lambda+2)x^3 - 2\lambda(\lambda+1)x \tag{22}$$

or in general (for odd n),

$$C_n^{\lambda}(x) = \sum_{\text{odd}\,\ell}^{n} G_{n,\ell} x^{\ell} \tag{23}$$

and the inverse which gives x^{ℓ} (ℓ odd) as a linear combination of Gegenbauer polynomials, that is,

$$x^{\ell} = \sum_{\text{odd}\,k}^{\ell} H_{\ell,k} C_k^{\lambda}(x). \tag{24}$$

We do not provide the explicit expressions for the matrices $G_{n,\ell}$ and $H_{\ell,k}$. These can be generated easily to all orders with the explicit expressions in Eq. (22) and the recurrence relation for the Gegenbauer polynomials given by

$$(\ell+2)C_{\ell+2}^{\lambda} = 2(\lambda+\ell+1)xC_{\ell+1}^{\lambda} - (2\lambda+\ell)C_{\ell}^{\lambda}.$$

With the substitution of Eq. (24) into Eq. (23), it is clear that

$$\sum_{\text{odd}\,\ell}^{n} G_{n,\ell} H_{\ell,k} = \delta_{n,k}. \tag{25}$$

We choose the very simple functions $f(x)=x$ and x^3 previously chosen [15,24] for which the exact Gegenbauer expansions are:

$$x = \frac{1}{2\lambda} C_1^{\lambda}(x),$$

$$x^3 = \frac{3}{4\lambda(\lambda+2)} C_1^{\lambda}(x) + \frac{3}{4\lambda(\lambda+1)(\lambda+2)} C_3^{\lambda}(x) \tag{26}$$

and the exact Fourier sine coefficients are given by the following:

$$b_k^{(1)} = \frac{2(-1)^{(k+1)}}{\pi k},$$

$$b_k^{(3)} = \frac{2(-1)^k}{\pi k}\left[\frac{6}{(\pi k)^2} - 1\right], \tag{27}$$

respectively. The optimal values of m in the sum over Gegenbauer polynomials, Eq. (6), is $m=1$ and 3, respectively, *irrespective* of the value of λ. If the sums are taken beyond these limits in each case, the approximation worsens and the sum, Eq. (6), approaches $f_N(x)$ containing the Gibbs phenomenon that is not desired. This represents a somewhat different viewpoint than the one by Gottlieb and co-workers based on their detailed numerical analysis where $\lambda = m = N/4$.

To demonstrate that the inverse method is exact for polynomials, we consider $f(x)=x^p$, odd p, with Fourier coefficients

$$b_k^{(p)} = \int_{-1}^{1} x^p \sin(k\pi x)\,\mathrm{d}x. \tag{28}$$

With Eq. (24), the exact Gegenbauer coefficients for this function are $g_\ell^{(p)} = H_{p,\ell}$. We set $\tilde{g}_\ell = H_{p,\ell}$ and show that Eq. (15) is an identity, that is,

$$\sum_{\text{odd}\,\ell}^{m} H_{p,\ell} T_{\ell,k} = b_k^{(p)}. \tag{29}$$

With the definition of $T_{\ell,k}$, Eq. (16), and the use of Eq. (23), we have that

$$\sum_{\text{odd}\,\ell}^{m} H_{p,\ell} \sum_{\text{odd}\,n}^{\ell} G_{\ell,n} b_k^{(n)} = b_k^{(p)}. \tag{30}$$

Finally, with Eq. (25) we get the result

$$\sum_{\text{odd}\,n}^{m} \delta_{p,n} b_k^{(n)} = b_k^{(p)} \tag{31}$$

and hence the relation, Eq. (15) provides an *exact* relationship between the first m (exact) $b_k^{(p)}$ Fourier coefficients and the first m $\tilde{g}_\ell = g_\ell$ Gegenbauer coefficients. Moreover, the results are independent of λ.

We now consider the functions employed previously [14,15,24] in the study of the direct method. Thus for the function $f(x) = x$, we need only one Fourier coefficient in Eq. (15) and

$$\tilde{g}_1 T_{1,1} = b_1^{(1)} \tag{32}$$

to get an *exact* resolution of the Gibbs phenomenon, that is $\tilde{g}_1 = g_1 = 1/2\lambda$. Similarly, for $f(x) = x^3$, we need solve the 2×2 system

$$\begin{aligned} \tilde{g}_1 T_{1,1} + \tilde{g}_3 T_{3,1} &= b_1^{(3)}, \\ \tilde{g}_1 T_{1,2} + \tilde{g}_3 T_{3,2} &= b_2^{(3)}. \end{aligned} \tag{33}$$

The $\tilde{g}_\ell = g_\ell$ Gegenbauer coefficients vary with λ (see Eq. (26)) but if Eq. (33) is written for the Taylor coefficients we have the system of equations:

$$\begin{aligned} d_1 b_1^{(1)} + d_3 b_1^{(3)} &= b_1^{(3)}, \\ d_1 b_2^{(1)} + d_3 b_2^{(3)} &= b_2^{(3)}, \end{aligned} \tag{34}$$

which clearly has the solution $d_1 = 0$ and $d_3 = 1$, independent of λ. We can write $\tilde{f}(x) = \sum_{\text{odd}\,n}^{m} d_n x^n$, where $d_n = \sum_{\text{odd}\,\ell}^{m} \tilde{g}_\ell G_{\ell,n}$, so that Eq. (15) for the Gegenbauer coefficients is transformed into the linear system,

$$\sum_{\text{odd}\,n}^{m} d_n b_k^{(n)} = b_k \tag{35}$$

for the Taylor coefficients which are independent of λ. If the b_k coefficients on the RHS of Eq. (35) are for $f(x) = x^p$ as given by Eq. (28), then Eq. (35) gives an exact result. One of the columns in the matrix $b_k^{(n)}$ is equal to the inhomogenous vector on the RHS of Eq. (35) and $d_n = 1$ for $n = p$ and $d_n = 0$ for $n \neq p$. This has been verified by the numerical solution of Eq. (35). An important

aspect of Eqs. (32)–(35) is that the exact b_k Fourier coefficients are considered as input. Therefore, the exact resolution of Gibbs phenomena requires that the exact $T_{k,\ell}$ matrix elements be computed.

5. Numerical examples

In this section, we consider several numerical examples some of which have been considered previously [15,24], and compare the inverse and direct methods. Our first choice is the polynomial $f(x) = x^3 + x^6$. The $\tilde{g}_\ell$ coefficients are obtained by inverting Eq. (15) and

$$\sum_{\text{even } \ell}^{m} \tilde{g}_\ell V_{\ell,k} = a_k, \tag{36}$$

where $V_{\ell,k}$ is given by Eq. (21) and $m=7$. We calculate the $T_{k,\ell}$ and $V_{k,\ell}$ matrix elements analytically as discussed in the appendix. The Taylor coefficients, d_n, determined from the $\hat{g}_\ell$ coefficients are shown in Table 1 versus λ. As can be seen from the table, the reconstruction of the function is exact to machine accuracy. Obviously the Gegenbauer coefficients vary with λ, but the representation of the function in the primitive basis functions is (essentially) independent of λ as is the L_∞ error, defined by

$$L_\infty = \max_{-1 \leqslant x \leqslant 1} |f(x) - \tilde{f}_m(x)|.$$

As might be expected the convergence of this inverse approach is extremely rapid and requires only 6 terms in Eqs. (15) and (36). By contrast, the convergence of the direct method, shown in Table 2 is considerably slower. Here, we show the convergence of the Taylor coefficients versus N in accordance with the reconstruction constraint in the direct method, that is, $m = \lambda = N/4$. As N increases, both m and λ increase. It is useful to note the decrease of the L_∞ error with increasing

Table 1
Inverse method; Taylor coefficients d_n versus λ for $f(x) = x^3 + x^6$

λ	d_0	d_1	d_2	d_3
$\frac{1}{2}$	0.132(−15)	0.000000	−0.278(−15)	1.0000
1	0.798(−16)	0.191(−15)	0.222(−15)	1.0000
2	−0.520(−17)	0.209(−15)	0.555(−15)	1.0000
4	0.183(−15)	0.393(−15)	−0.611(−15)	1.0000
8	−0.316(−15)	0.135(−15)	0.125(−14)	1.0000

	d_4	d_5	d_6	d_7	L_∞
$\frac{1}{2}$	0.222(−15)	0.000000	1.0000	0.000000	0.444(−15)
1	0.444(−15)	0.556(−14)	1.0000	−0.389(−14)	0.257(−15)
2	−0.155(−14)	0.556(−14)	1.0000	−0.389(−14)	0.444(−15)
4	0.555(−15)	0.103(−13)	1.0000	−0.779(−14)	0.133(−14)
8	0.111(−15)	0.860(−15)	1.0000	0.000000	0.666(−15)

$(-n) \equiv 10^{-n}$; $m = 7$.

Table 2
Direct method; convergence of Taylor coefficients d_n versus N for $f(x) = x^3 + x^6$

N	d_0	d_1	d_2	d_3	
4	0.802(−1)	0.462(0)			
8	−0.314(−1)	0.377(0)	0.376(0)		
12	−0.179(−1)	−0.677(−3)	0.268(0)	1.002652	
16	0.372(−2)	0.180(−4)	−0.156(0)	0.999909	
20	0.260(−2)	−0.134(−4)	−0.125(0)	1.000160	
24	−0.111(−7)	0.601(−7)	0.669(−6)	0.999999	
28	−0.809(−10)	−0.257(−6)	0.457(−8)	1.000010	
32	−0.120(−9)	−0.513(−8)	0.129(−7)	1.000000	
36	−0.897(−12)	−0.461(−8)	0.394(−9)	1.000000	
40	0.135(−11)	−0.243(−9)	0.219(−9)	1.000000	

	d_4	d_5	d_6	d_7	L_∞
4					0.146(1)
8					0.128(1)
12					0.752(0)
16	0.833(0)				0.319(0)
20	0.750(0)	−0.317(−3)			0.373(0)
24	−0.490(−5)	0.305(−5)	1.000010		0.559(−5)
28	−0.292(−7)	−0.322(−4)	1.000000	0.430(−4)	0.167(−4)
32	−0.181(−6)	−0.642(−6)	1.000000	0.844(−6)	0.595(−6)
36	−0.587(−8)	−0.180(−5)	1.000000	0.586(−5)	0.166(−5)
40	−0.525(−8)	−0.964(−7)	1.000000	0.322(−6)	0.117(−6)

$m = \lambda = N/4$.

N, and the comparison for $N = 40$ with the value in Table 1. The role of the parameter λ will become clearer in the discussion to follow. Here, we note that for this function the optimal choice of m is $m = 6$, whereas in the direct method $m = 10$ for $N = 40$.

To understand the role of λ in the direct method, it is useful to consider the matrices $\mathbf{U}^{(s)}$ and $\mathbf{U}^{(c)}$, given by Eqs. (18) and (20). As noted earlier, these are unit matrices only in the limit $N \to \infty$. To illustrate the departure of these matrices from unit matrices we show in Fig. 1 the variation of $\mathrm{tr}(\mathbf{I} - \mathbf{U}^{(s)})$ and $\mathrm{tr}(\mathbf{I} - \mathbf{U}^{(c)})$ versus λ, where $\mathbf{I}$ is the unit matrix. As can be seen from the figure, these functions are zero for specific values of λ. These roots occur for values of λ greater than about 2 which accounts for the poor performance of Chebyshev polynomials ($\lambda = 0$) and Legendre polynomials ($\lambda = 1/2$) with the direct method. With an increase in λ, the amplitudes of the oscillations in the functions in Fig. 1 diminish and hence the reconstruction of the function is expected to be better for larger λ values. In Table 3, we show the variation of the Taylor coefficients versus λ and they either tend to zero or oscillate about unity. The underlined and bold-faced entries in the table indicate where d_3 and d_6 cross unity at values of λ that correspond to the roots of the functions in Fig. 1 with $N = 32$ and $m = N/4 = 8$.

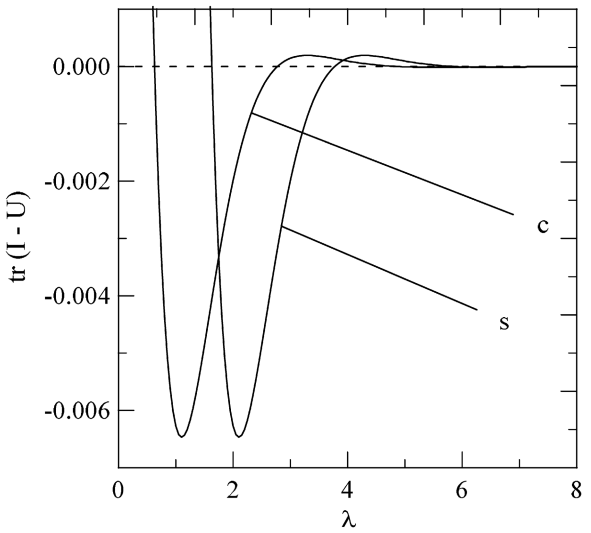

Fig. 1. Variation of $\mathrm{tr}(\mathbf{I}-\mathbf{U}^{(s)})$ and $\mathrm{tr}(\mathbf{I}-\mathbf{U}^{(c)})$ vs. λ for $N=76$ and $m=19$; see Eqs. (18) and (20).

The second function that is chosen is $f(x)=\cos[1.4\pi(x+1)]$. The Fourier coefficients are explicitly:

$$a_0=\frac{\sin(2.8\pi)}{(2.8\pi)},$$

$$a_k=\frac{1.4\pi(-1)^{(k+1)}}{(k\pi)^2}\,\frac{\sin(2.8\pi)}{(1-(\frac{1.4}{k})^2)},$$

$$b_k=\frac{k\pi(-1)^k}{(1.4\pi)^2}\,\frac{(\cos(2.8\pi)-1)}{(1-(\frac{k}{1.4})^2)}.$$

The results with the inverse method are shown in Table 4 which gives the Taylor coefficients versus λ, and as with the previous example, the results are independent of λ. The L_∞ error is also shown and is close to machine accuracy for $m=23$ and $N=12$. With only 23 Gegenbauer polynomials, it is clear that the higher-order Taylor coefficient d_{19} is unconverged. The d_{19} coefficient differs from the exact value by 0.53%. These results illustrate that the resolution of the Gibbs phenomenon ultimately depends on the rate of convergence of the power series of the functions considered.

From the results presented (see Table 2) it is clear that if m is not sufficiently large (say $m<6$, $N<24$, for $f(x)=x^3+x^6$) then the reconstructed function with the direct method is inaccurate. On the other hand if m is too large, then the reconstructed function will exhibit some of the Gibbs phenomena, since $\lim_{m\to\infty}\hat{f}_m(x)=f_N(x)$. In the application of the inverse method, a criterion is required to truncate the sum over m and here we consider the residue $R(m)$ defined by

$$R(m)=\left|\sum_{i=0}^{m}|\tilde{g}_i^{\lambda}|-\sum_{i=0}^{m-1}|\tilde{g}_i^{\lambda}|\right|, \tag{37}$$

Table 3
Direct method; Taylor coefficients d_n versus λ for $f(x) = x^3 + x^6$

λ/d_n	d_0	d_1	d_2	d_3	d_4	d_5	d_6	d_7
0.5	−0.127(−3)	0.121(0)	0.536(−2)	−0.329716	−0.332(−1)	0.346(01)	1.06314	−0.248(01)
0.66838	−0.113(−4)	0.708(−1)	0.343(−3)	0.197608	−0.114(−2)	0.214(1)	**1.00000**	−0.156(1)
0.9	0.444(−4)	0.321(−1)	−0.219(−2)	0.623311	0.158(−1)	0.103(01)	0.965477	−0.774(0)
1.3	0.413(−4)	0.607(−2)	−0.211(−2)	0.925030	0.155(−1)	0.214(0)	0.965224	−0.166(0)
1.67753	0.207(−4)	0.560(−4)	−0.111(−2)	**1.00000**	0.853(−2)	−0.229(−2)	0.980236	0.378(−2)
1.7	0.197(−4)	−0.768(−4)	−0.106(−2)	1.00174	0.816(−2)	−0.754(−2)	0.981074	0.804(−2)
2.1	0.677(−5)	−0.827(−3)	−0.381(−3)	1.01197	0.306(−2)	−0.395(−1)	0.992643	0.349(−1)
2.5	0.146(−5)	−0.522(−3)	−0.852(−4)	1.00787	0.702(−3)	−0.269(−1)	0.998279	0.244(−1)
2.7	0.397(−6)	−0.353(−3)	−0.225(−4)	1.00544	0.179(−3)	−0.189(−1)	0.999580	0.175(−1)
2.80499	0.753(−7)	−0.279(−3)	−0.313(−5)	1.00434	0.138(−4)	−0.153(−1)	**1.00000**	0.142(−1)
2.9	−0.117(−6)	−0.221(−3)	0.866(−5)	1.00348	−0.885(−4)	−0.123(−1)	1.00026	0.116(−1)
3.1	−0.313(−6)	−0.127(−3)	0.211(−4)	1.00204	−0.199(−3)	−0.736(−2)	1.00055	0.701(−2)
3.5	−0.299(−6)	−0.284(−4)	0.205(−4)	1.00047	−0.197(−3)	−0.172(−2)	1.00056	0.166(−2)
3.7	−0.230(−6)	−0.751(−5)	0.161(−4)	1.00012	−0.157(−3)	−0.421(−3)	1.00045	0.384(−3)
3.81584	−0.190(−6)	−0.514(−6)	0.134(−4)	**1.00000**	−0.132(−3)	0.352(−4)	1.00038	−0.724(−4)
3.9	−0.162(−6)	0.289(−5)	0.115(−4)	0.999941	−0.114(−3)	0.264(−3)	1.00033	−0.305(−3)
4.1	−0.105(−6)	0.703(−5)	0.762(−5)	0.999867	−0.765(−4)	0.555(−3)	1.00022	−0.606(−3)
4.3	−0.627(−7)	0.775(−5)	0.461(−5)	0.999853	−0.470(−4)	0.616(−3)	1.00014	−0.674(−3)
4.5	−0.333(−7)	0.685(−5)	0.248(−5)	0.999868	−0.255(−4)	0.557(−3)	1.00008	−0.616(−3)
4.7	−0.145(−7)	0.537(−5)	0.108(−5)	0.999895	−0.111(−4)	0.449(−3)	1.00003	−0.502(−3)
4.964	−0.108(−8)	0.340(−5)	0.577(−7)	0.999932	−0.318(−6)	0.295(−3)	**1.00000**	−0.335(−3)
5.1	0.239(−8)	0.254(−5)	−0.215(−6)	0.999949	0.264(−5)	0.225(−3)	0.999991	−0.258(−3)
5.3	0.484(−8)	0.154(−5)	−0.412(−6)	0.999969	0.482(−5)	0.140(−3)	0.999984	−0.162(−3)
5.5	0.532(−8)	0.829(−6)	−0.455(−6)	0.999983	0.532(−5)	0.767(−4)	0.999982	−0.892(−4)
5.7	0.479(−8)	0.361(−6)	−0.414(−6)	0.999993	0.489(−5)	0.333(−4)	0.999983	−0.386(−4)
5.979	0.345(−8)	0.728(−8)	−0.303(−6)	**1.00000**	0.364(−5)	−0.119(−5)	0.999987	0.267(−5)
6.1	0.284(−8)	−0.718(−7)	−0.252(−6)	1.00000	0.304(−5)	−0.929(−5)	0.999989	0.126(−4)
6.3	0.194(−8)	−0.138(−6)	−0.174(−6)	1.00000	0.212(−5)	−0.163(−4)	0.999993	0.213(−4)

$N = 32, m = N/4 = 8$.

Table 4
Inverse method; variation of Taylor coefficients d_n versus λ for $f(x) = \cos[1.4\pi(x+1)]$

λ/d_n	d_0	d_5	d_{10}	d_{15}	d_{19}	L_∞
$\frac{1}{2}$	−0.309017	0.130441(2)	0.230674	−0.324243(−2)	−0.129876(−4)	0.7430(−12)
1	−0.309017	0.130441(2)	0.230674	−0.324245(−2)	−0.130249(−4)	0.1628(−12)
2	−0.309017	0.130441(2)	0.230674	−0.324244(−2)	−0.130052(−4)	0.2451(−12)
4	−0.309017	0.130441(2)	0.230674	−0.324250(−2)	−0.130888(−4)	0.1061(−11)
8	−0.309017	0.130441(2)	0.230674	−0.324241(−2)	−0.129744(−4)	0.1671(−11)
Exact	−0.309017	0.130441(2)	0.230674	−0.324247(−2)	−0.130434(−4)	

$N = 12,\ m = 23$.

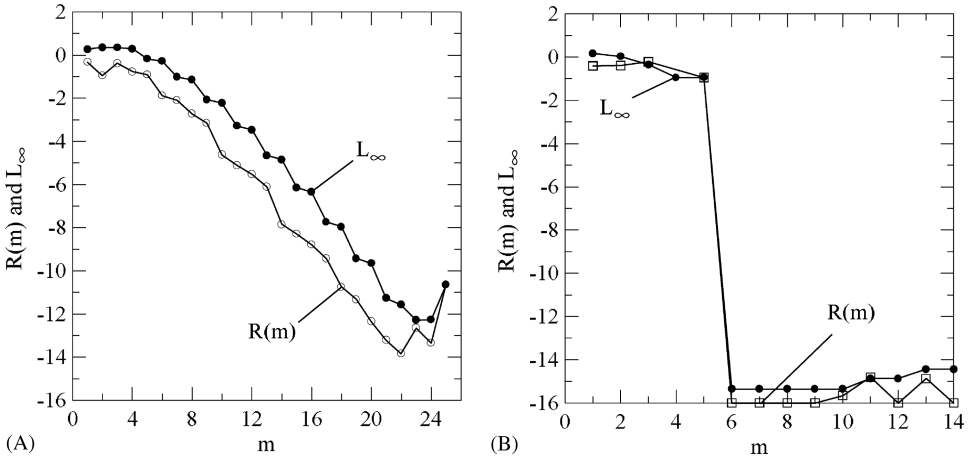

Fig. 2. The variation of the residue, $R(m)$ (squares), and the L_∞ error (circles) versus m. (A) $f(x)=\cos[1.4\pi(x+1)]$; (B) $f(x)=x^3+x^6$; results are independent of $\lambda=1$.

where $\tilde{g}_i^\lambda$ are the Gegenbauer coefficients that results form the inversion of m equations, and are not necessarily the same as $\tilde{g}_i^\lambda$ for $m+1$ equations. In Fig. 2A, we show the variation of the L_∞ error (solid circles) and the residue, $R(m)$ (open circles), versus m for the reconstruction of $f(x)=\cos[1.4\pi(x+1)]$. It is clear from the figure that an optimal value of m is approximately 22. Fig. 2B illustrates the same thing for the function $f(x)=x^3+x^6$ and the optimal value of m based on the residue is 6. In the direct method, the choice $m=N/4$ is made.

A comparison of the convergence of the direct and inverse methods for $f(x)=\cos[1.4\pi(x+1)]$ is shown in Fig. 3, and Tables 4 and 6. For the direct method $m=\lambda=N/4$, whereas for the inverse method $m=N$ and independent of λ. It is clear that the inverse method provides a much faster convergence than the direct method. It is of considerable interest to understand the role of λ in the direct method. We have already demonstrated that large values of λ give more accurate results in terms of the closeness of the matrices $U^{(s)}_{\ell,\ell'}$ and $U^{(c)}_{\ell,\ell'}$ to unit matrices (see Fig. 1). To get a better understanding of this behaviour, we have considered alternate weight functions in addition to the Gegenbauer weight function, $w_g(x)=(1-x^2)^{\lambda-1/2}$. These other weight functions are a modified Gegenbauer weight function, $w_m(x)=w_g(x)\mathrm{e}^{-\alpha x^2}$ and a gaussian weight function $w(x)=\mathrm{e}^{-\beta x^2}$. The parameters λ, α and β in the weight function control the width of the weight functions about the origin. The orthogonal polynomials are constructed with the Gautschi Stieltjes procedure [10] as discussed elsewhere [5,7,16,20]. For the Gegenbauer and modified Gegenbauer weight functions, $m=\lambda=N/4$. It is clear that the inverse method converges faster than the direct method with different polynomial basis sets (Fig. 3B).

The variation of $\mathrm{tr}(\mathbf{I}-\mathbf{U}^{(s)})$ and $\mathrm{tr}(\mathbf{I}-\mathbf{U}^{(c)})$ versus λ are shown in Figs. 4A and 4B for $N=44$ and $m=25$, and $N=32$ and $m=17$, respectively. The roots vary with N, m and λ. The behaviour in Fig. 4 is useful in understanding the variation of the L_∞ error versus λ shown in Fig. 5A. The oscillations of the error versus λ for different N and m follow the oscillations shown in Fig. 4 and the minima in Fig. 5 correspond to the roots in Fig. 4. In Fig. 5B, we show the L_∞ error versus N for several fixed values of λ and also for $\lambda=N/2$ and $\lambda=N/4$. It is clear that the smaller fixed

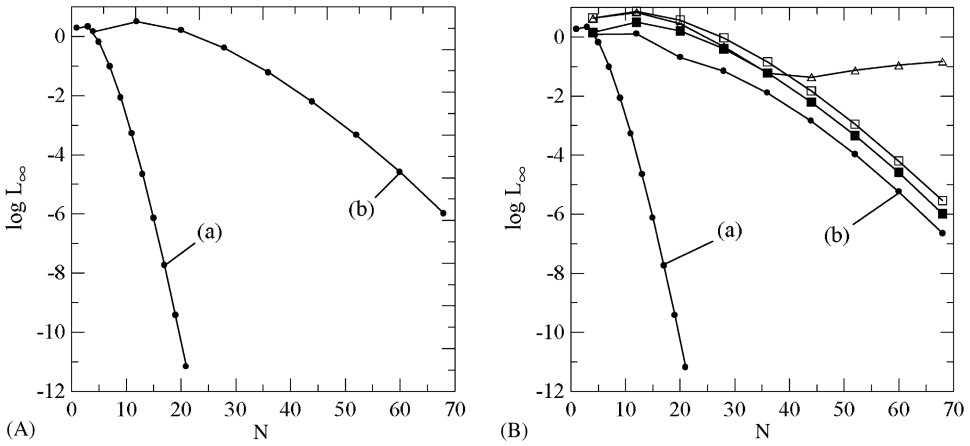

Fig. 3. The variation of the L_∞ error versus N for $f(x)=\cos[1.4\pi(x+1)]$; comparison of the direct and inverse methods. (A) Comparison of the inverse method and the direct method based on the Gegenbauer weight, $w_g(x)=(1-x^2)^{\lambda-1/2}$; (a) inverse method, (b) direct method; (B) comparison of the inverse method and the direct method based on different weight functions: Gegenbauer weight function, $\lambda=N/4$ (filled squares); modified Gegenbauer weight, $w_m(x)=w_g(x)\mathrm{e}^{-\alpha x^2}$, $\lambda=N/4$ (open squares), $\alpha=16$; filled circles, $\alpha=-16$; and a gaussian weight function $w(x)=\mathrm{e}^{-\beta x^2}$, $\beta=16$ (triangles).

Table 5
Direct method; variation of $d_n(\lambda)$ versus λ for $f(x)=\cos[1.4\pi(x+1)]$

λ	d_0	d_5	d_{10}	d_{15}	d_{19}	L_∞
$\frac{1}{2}$	−0.309029	−0.203631(3)	0.151266(2)	0.454691(5)	0.566650(4)	0.5352
1	−0.309016	−0.381491(2)	−0.175522(1)	0.123990(5)	0.161670(4)	0.2472
2	−0.309017	0.133678(2)	−0.614139	−0.143695(3)	−0.228711(2)	0.2152(−1)
4	−0.309017	0.130498(2)	0.242121	−0.258823(1)	−0.402494	0.1358(−3)
6	−0.309017	0.130439(2)	0.230461	0.166745	0.336778(−1)	0.1112(−3)
8	−0.309017	0.130441(2)	0.230678	−0.118250(−1)	−0.203508(−2)	0.1500(−4)
10	−0.309017	0.130441(2)	0.230674	−0.279070(−2)	0.112471(−3)	0.1613(−5)
14	−0.309017	0.130441(2)	0.230674	−0.326321(−2)	−0.173321(−4)	0.6925(−7)
18	−0.309017	0.130441(2)	0.230674	−0.325798(−2)	−0.181841(−4)	0.2983(−6)
Exact	−0.309017	0.130441(2)	0.230674	−0.324247(−2)	−0.130434(−4)	

$N=76, m=N/4=19$.

values of λ ((1) and (5)) do not give accurate results whereas the larger values (11 and 16) give results comparable to the choice $\lambda=N/4$ (Fig. 5b, Tables 5 and 6).

The numerical calculations considered in the present work with the inverse method involved the inversion of Eqs. (15) and (36) for the $\tilde{g}_\ell$ coefficients and the reconstructed function, Eq. (13). As has been demonstrated for the limited test functions studied in this section, the method gives remarkable results. Alternatively, the $\tilde{g}_\ell$ coefficients can also be determined from the inversion of Eq. (17) or Eq. (36) for the equivalent Taylor coefficients. One concern is the numerical invertibility of the matrices, $T_{k,\ell}$, $U_{\ell,\ell'}, V_{k,\ell}$ and $b_k^{(n)}$. Although the condition numbers of these matrices vary differently

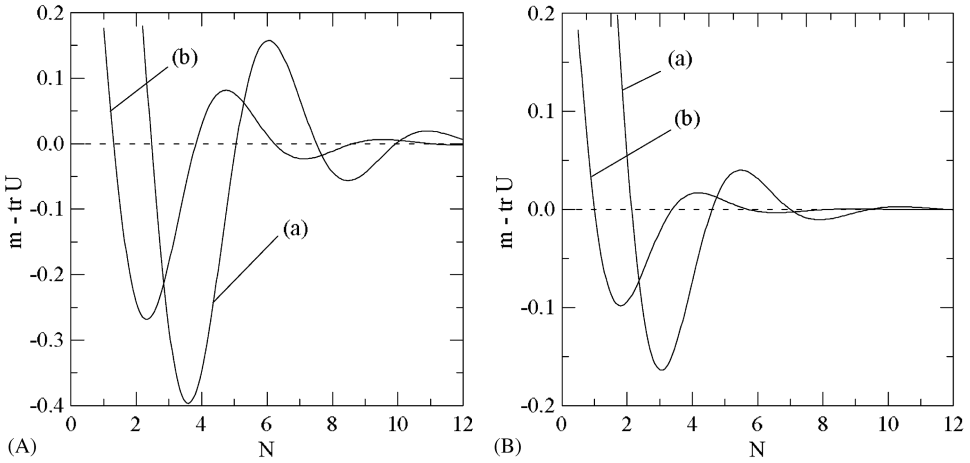

Fig. 4. Variation of (a) $\mathrm{tr}(\mathbf{I}-\mathbf{U}^{(s)})$ and (b) $\mathrm{tr}(\mathbf{I}-\mathbf{U}^{(c)})$ vs. λ (A) $N=44$ and $m=25$; (B) $N=32$ and $m=17$.

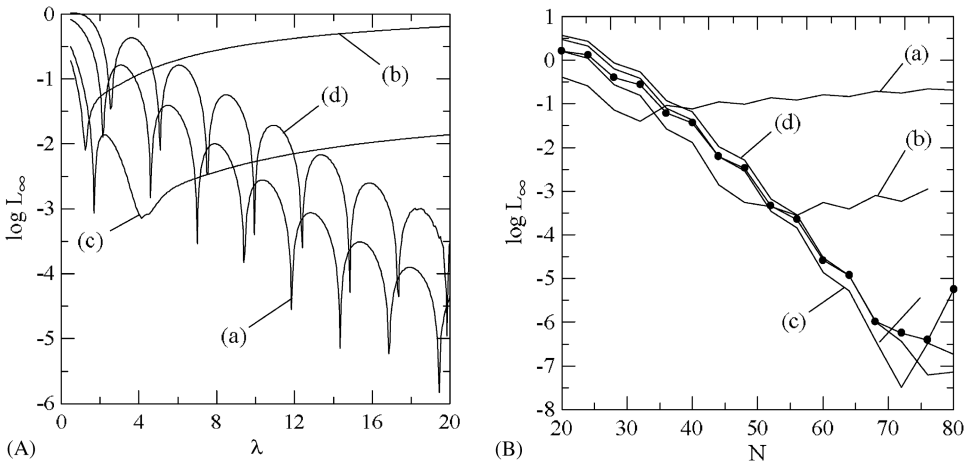

Fig. 5. Direct method: Variation of L_∞ for $f(x)=\cos[1.4\pi(x+1)]$, (A) vs. λ; N and m are equal to (a) 32, 17, (b) 32, 8, (c) 44, 11, and (d) 44, 25; (B) versus $N=4m$ with λ equal to (a) 1, (b) 5, (c) 11, and (d) 16 (filled circles), $\lambda=N/4$.

with m, it was found that an accurate resolution of the Gibbs phenomenon, $L_\infty \leqslant 10^{-8}$, can be obtained for most functions that can be approximated by a power series of order less than about 25. This includes a very large class of functions. A more detailed investigation of the invertibility of the these matrices is in progress and will be reported in a subsequent paper.

We consider an additional test function employed by Driscoll and Fornberg [8] given by, $f(x)=\exp[\sin(2.7x)+\cos(x)]$. In Fig. 6, we show the results of the direct (solid curves) and inverse (dotted curves) methods in comparison with the results reported in Fig. 7(d) of Ref. [8] (dashed curves). In Ref. [8], the oscillations in the error was not shown and a continuous curve connecting the maxima of the error versus x was drawn. The results obtained with the inverse method are clearly superior to either the singular Fourier–Padé method or the direct method. The singular Fourier–Padé

Table 6
Direct method; convergence of $d_n(\lambda)$ versus N for $f(x) = \cos[1.4\pi(x+1)]$

N	d_0	d_5	d_{10}	d_{15}	d_{19}	L_∞
24	−0.309017	0.128877(2)	0.119541(1)	0.231045(3)	0.536584(2)	0.3068
34	−0.309017	0.130492(2)	0.231432	−0.855467(1)	−0.213193(1)	0.1750(−1)
44	−0.309017	0.130441(2)	0.230647	0.358826(−1)	0.112343(−1)	0.1514(−3)
54	−0.309017	0.130441(2)	0.230674	−0.317325(−2)	0.135262(−4)	0.7288(−6)
64	−0.309017	0.130441(2)	0.230674	−0.324241(−2)	−0.102702(−4)	0.4081(−7)
74	−0.309017	0.130441(2)	0.230674	−0.326022(−2)	−0.207594(−4)	0.1801(−6)
Exact	−0.309017	0.130441(2)	0.230674	−0.324247(−2)	−0.130434(−4)	

$m = 19, \lambda = N/4$.

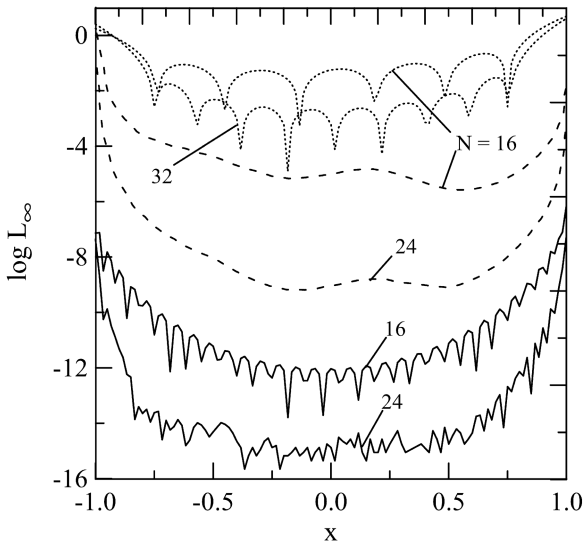

Fig. 6. Reconstruction of $f(x) = \exp[\sin(2.7x) + \cos(x)]$. The upper dotted curves are for the direct method. The dashed curves are for the singular Fourier–Padé method taken from Fig. 7(d) of Ref. [8]. The solid curves are the present results with the inverse method.

method does provide better results than the direct method. It is important to note that the singular Fourier–Padé method does not provide exact results for polynomials as does the inverse method (Fig. 6).

6. Local reconstruction by the inverse method

In this section, we consider the case that the function $f(x)$ is only piecewise continuous for a given interval and show how the inverse method described in the previous section in a single interval can be implemented for the reconstruction of $f(x)$ in multi-intervals. We assume that the

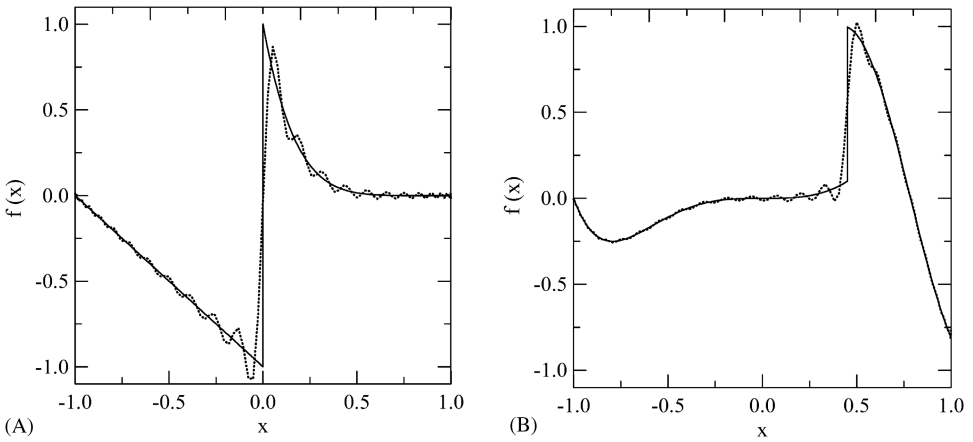

Fig. 7. (A) Reconstruction of $f(x)=-1-x$ if $x \leqslant 0$ and $(1-x)^6$ if $x > 0$. Oscillatory curve is the Chebyshev representation with $N = 35$. (B) $f(x)=x^3+x^6$ if $x \leqslant 0.45$ and $\cos[1.4\pi(x+1)]$. The solid lines are the reconstructed function from the inversion of Eq. (46) with $m_R = 8$ and $m_L = 12$, and indistinguishable from the original function (see Tables 7 and 8).

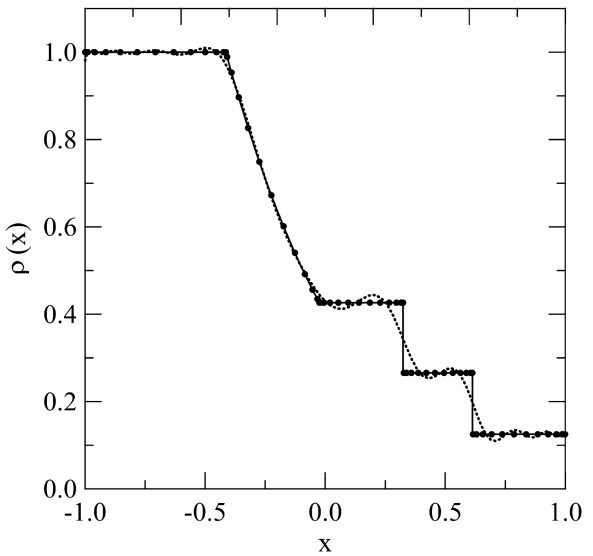

Fig. 8. Reconstruction of the density profile for Sod's problem. Oscillatory curve is the Chebyshev representation with $N = 23$. The solid line is the reconstructed function from the inversion of Eq. (44) or Eq. (42), and indistinguishable from the original function.

function $f(x)$ is piecewise analytic and that all the locations where either $f(x)$ or its first derivative are discontinuous are known. The inverse method, as does the direct method, recovers the function $f(x)$ piecewise in each subinterval to remove the Gibbs oscillations around a discontinuity. A major aspect of this resolution procedure is to identify the location of discontinuities. This is not addressed in the present paper and we note that there are edge detection methods available [2,11]. Here, we simply assume that these are known.

Suppose that $f(x)$ is piecewise analytic defined in the interval $x \in \Omega \equiv [-1,1]$ and that there are two subintervals of unequal length, one on the left and one on the right denoted by $\Omega^{\mathrm{L}} = [-1,a]$ and $\Omega^{\mathrm{R}} = [a,1]$, where $-1 < a < 1$. The generalization to more than two intervals is straightforward. We assume that $f(x)$ is analytic in each subinterval. Let $f^{\mathrm{L}}(x)$ and $f^{\mathrm{R}}(x)$ denote $f(x)$ in Ω^{L} and Ω^{R}, respectively. Hereafter, we will refer to L and R in superscript or subscript as the left interval and the right, respectively.

Since $f^{\mathrm{L}}(x)$ and $f^{\mathrm{R}}(x)$ are assumed to be analytic we want to represent them as the sum of the Gegenbauer polynomials such that

$$f^{\mathrm{L}}(y) = \sum_{l=0}^{\infty} g_l^{\mathrm{L}} C_l^{\lambda}(X_{\mathrm{L}}(y)), \quad y \in \Omega^{\mathrm{L}},$$

$$f^{\mathrm{R}}(y) = \sum_{l=0}^{\infty} g_l^{\mathrm{R}} C_l^{\lambda}(X_{\mathrm{R}}(y)), \quad y \in \Omega^{\mathrm{R}}, \tag{38}$$

where

$$X_{\mathrm{L}}(y) = \frac{2y+1-a}{1+a}$$

and

$$X_{\mathrm{R}}(y) = \frac{2y-1-a}{1-a},$$

such that $-1 \leqslant X_{\mathrm{L}}, X_{\mathrm{R}} \leqslant 1$. As in the previous section we do not know the function $f(x)$ but we are given the expansion of $f(x)$ in a set of orthogonal functions $T_k(x)$ that satisfy an othogonality relation

$$\int_{-1}^{1} w(x) T_k(x) T_{\ell}(x)\,\mathrm{d}x = c_k \delta_{k,\ell}, \tag{39}$$

where $w(x)$ is the weight function. In the previous sections, these were the Fourier sine and cosine functions, but these could also be the Chebyshev, Legendre or other polynomial basis sets. In this work, we choose the Chebyshev polynomials which are routinely used by the spectral community, and we thus have the representation,

$$f_N(x) = \sum_{i=0}^{N} b_k T_k(x). \tag{40}$$

The function $f_N(x)$ will exhibit Gibbs oscillations around a discontinuity. With Eqs. (38) and (39), the expansion coefficients b_k are given by

$$\begin{aligned} b_k &= \frac{1}{c_k} \int_{-1}^{a} w(x) f^{\mathrm{L}}(x) T_k(x)\,\mathrm{d}x + \frac{1}{c_k} \int_{a}^{1} w(x) f^{\mathrm{R}}(x) T_k(x)\,\mathrm{d}x \\ &= \sum_{l=0}^{\infty} g_l^{\mathrm{L}} \frac{1}{c_k} \int_{-1}^{a} w(y) C_l^{\lambda}[X_{\mathrm{L}}(y)] T_k(y)\,\mathrm{d}y + \sum_{l=0}^{\infty} g_l^{\mathrm{R}} \frac{1}{c_k} \int_{a}^{1} w(y) C_l^{\lambda}[X_{\mathrm{R}}(y)] T_k(y)\,\mathrm{d}y. \end{aligned} \tag{41}$$

We now transform the integrations over y in each subinterval to integrals over X_{L} and X_{R}, so that

$$b_k = \sum_{l=0}^{\infty} g_l^{\mathrm{L}} \frac{(1+a)}{2c_k} \int_{-1}^{1} w[y(X_{\mathrm{L}})] C_l^{\lambda}(X_{\mathrm{L}}) T_k[y(X_{\mathrm{L}})]\, \mathrm{d}X_{\mathrm{L}} + \sum_{l=0}^{\infty} g_l^{\mathrm{R}} \frac{(1-a)}{2c_k} \int_{-1}^{1} w[y(X_{\mathrm{R}})] C_l^{\lambda}(X_{\mathrm{R}}) T_k[y(X_{\mathrm{R}})]\, \mathrm{d}X_{\mathrm{R}}, \tag{42}$$

We identify the matrices $L_{k,\ell}$ and $R_{k,\ell}$ in Eq. (42), defined by,

$$L_{k,\ell} = \frac{(1+a)}{2c_k} \int_{-1}^{1} w[y(X_{\mathrm{L}})] C_l^{\lambda}(X_{\mathrm{L}}) T_k[y(X_{\mathrm{L}})]\, \mathrm{d}X_{\mathrm{L}},$$

$$R_{k,\ell} = \frac{(1-a)}{2c_k} \int_{-1}^{1} w(y(X_{\mathrm{R}}) C_l^{\lambda}(X_{\mathrm{R}}) T_k[y(X_{\mathrm{R}})]\, \mathrm{d}X_{\mathrm{R}}. \tag{43}$$

Thus the N b_k coefficients are given by

$$b_k = \sum_{l=0}^{\infty} L_{k,\ell} g_l^{\mathrm{L}} + \sum_{l=0}^{\infty} R_{k,\ell} g_l^{\mathrm{R}}, \quad k = 0, \ldots, N. \tag{44}$$

Now we want to approximate g_l^{L} and g_l^{R} as $\tilde{g}_l^{\mathrm{L}}$ and $\tilde{g}_l^{\mathrm{R}}$ by truncating the above infinite series by m_{L} and m_{R} for Ω^{L} and Ω^{R}, respectively, such that

$$b_k = \sum_{l=0}^{m_{\mathrm{L}}} L_{k,\ell} \tilde{g}_l^{\mathrm{L}} + \sum_{l=0}^{m_{\mathrm{R}}} R_{k,\ell} \tilde{g}_l^{\mathrm{R}}, \quad k = 0, \ldots, N. \tag{45}$$

It is important to note that m_{L} is not necessarily equal to m_{R}. Eq. (45) is the generalization of Eq. (15) to two intervals. In order to invert Eq. (45) for the two sets of unknowns, $\tilde{g}_\ell^{\mathrm{R}}$ and $\tilde{g}_\ell^{\mathrm{L}}$ we introduce a vector notation such that $\mathbf{b} = (b_0, \ldots, b_N)^{\mathrm{T}}$ and $\tilde{\mathbf{g}} = (\tilde{g}_0^{\mathrm{L}}, \ldots, \tilde{g}_{m_{\mathrm{L}}}^{\mathrm{L}}, \tilde{g}_0^{\mathrm{R}}, \ldots, \tilde{g}_{m_{\mathrm{R}}}^{\mathrm{R}})^{\mathrm{T}}$ with $m_{\mathrm{L}} + m_{\mathrm{R}} + 2$ elements. It is important to note that the matrices $\mathbf{L}$ and $\mathbf{R}$ are not square matrices and are of dimension $(N+1)\times(m_{\mathrm{L}}+1)$ and $(N+1)\times(m_{\mathrm{R}}+1)$, respectively. If we set $m \equiv m_{\mathrm{L}}+m_{\mathrm{R}}=N-1$, we can construct from these matrices the square $(N+1)\times(N+1)$ matrix $\mathbf{T}=[\mathbf{L}\,|\,\mathbf{R}]$ which explicitly is given by

$$\mathbf{T} = \begin{bmatrix} L_{0,0} & \cdots\cdots & L_{0,m_{\mathrm{L}}}, & R_{0,0} & \cdots\cdots & R_{0,m_{\mathrm{R}}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ L_{N,0} & \cdots\cdots & L_{N,m_{\mathrm{L}}}, & R_{N,0} & \cdots\cdots & R_{N,m_{\mathrm{R}}} \end{bmatrix}. \tag{46}$$

Then Eq. (45) can be rewritten as

$$\mathbf{b} = \mathbf{T} \cdot \tilde{\mathbf{g}}. \tag{47}$$

Table 7
Inverse method; L_∞ error for two sub-domains

m_L	m_R	N	L_∞ (left)	L_∞ (right)
1	6	8	0.666134(−15)	0.121347(−12)
2	7	10	0.999201(−15)	0.224043(−12)
3	9	13	0.296430(−13)	0.248157(−11)
5	5	11	0.444679(−2)	0.100234(−1)

$f(x) = -1 - x$ if $x \leqslant 0$ (left), and $f(x) = (1-x)^6$ if $x > 0$ (right).

Table 8
Inverse method; L_∞ error for two sub-domains

m_L	m_R	N	L_∞ (left)	L_∞ (right)
6	8	15	0.481193(−10)	0.563992(−6)
6	9	16	0.763695(−14)	0.791313(−8)
6	10	17	0.164141(−12)	0.444790(−9)
6	11	18	0.123716(−12)	0.128884(−7)

$f(x) = x^3 + x^6$ if $x \leqslant 0.45$ (left), and $f(x) = \cos[1.4\pi(x+1)]$ if $x > 0.45$ (right).

It is interesting to observe that it is not necessary to use all given modes up to N in order to reconstruct the given function. If both $f^L(x)$ and $f^R(x)$ belong to a polynomial family of degree at most m_L and m_R, respectively, then $m = m_L + m_R < N - 1$ is sufficient.

To show how the inverse method works for the local reconstructions in multi-intervals three numerical examples are provided. For the first numerical example, we consider a polynomial in each interval, that is, $f(x) = -1 - x$ if $x \leqslant 0$ and $(1-x)^6$ if $x > 0$. Chebychev polynomials are used and the calculations of the matrix $\mathbf{T}$ and the b_k coefficients are carried out with 600 quadrature points for each subinterval. Table 7 shows the L_∞ error for different choice of (m_L, m_R) for the first example. In Table 7 we find that the reconstruction is exact to within machine accuracy if $m_L = 1$ and $m_R = 6$ as the theory predicts, and that $m = N - 1 = 7$ is sufficient. The second example sets $f(x) = x^3 + x^6$ if $x \leqslant 0.45$ and $\cos[1.4\pi(x+1)]$ if $0.45 < x \leqslant 1$. In Table 8 the L_∞ errors for different (m_L, m_R) are given. Table 8 shows the L_∞ error for $(m_L, m_R) = (6, 9)$ reaches machine accuracy consistent with the choice of a sixth order polynomial in the left domain. In Fig. 6, $f_N(x)$ is represented as a dotted line and the reconstruction as a solid line. We use $(m_L, m_R) = 1, 6$ and $(m_L, m_R) = 6, 9$ and $N = 8$ and 16 for Figs. 7(a) and 7(b), respectively. As can be seen from the figures, the local inverse method gives excellent results. It is important to point out that the present inverse method will give exact results for the test functions, $f_b(x) = |x|/\pi$ and $f_c(x) = \max(0, x)/\pi$ considered by Driscoll and Fornberg [8], whereas their singular Fourier–Padé method does not provide exact results (see Figs. 5b and 5c of Ref. [8]).

Solutions to time-dependent nonlinear hyperbolic compressible hydrodynamic equations are often characterized by singularities or sharp discontinuities even for the smooth initial conditions. The

time dependent solutions can become nonsmooth and singular, and conservative numerical schemes can suffer from severe Gibbs oscillations around the discontinuities, and the overall accuracy of the scheme is destroyed or sometimes the scheme goes quickly unstable [12]. Thus, an important aspect of numerical schemes adapted for such systems is to incorporate a shock capturing methodology. When one considers a numerical calculation of such fluid dynamical systems, Sod's shock tube problem [21,6,25] is widely used as a test problem to check a particular shock capturing scheme. The one-dimensional Sod shock problem is described by the 1D Euler gas equations, that is

$$\frac{\partial \mathbf{q}(x,t)}{\partial t} = \frac{\partial f(\mathbf{q})}{\partial x}.$$

The vector $\mathbf{q}(x,t) = (\rho, \rho u, E)^{\mathrm{T}}$ and the flux vector $f(q,t) = (\rho u, P + \rho u^2, (E+P)u)^{\mathrm{T}}$, where ρ, u, E and P are the dimensionless density, gas velocity, total energy and pressure, respectively. The initial condition is given by

$$q(x,0) = \begin{cases} (1,0,2.5)^{\mathrm{T}} & \text{if } x \leqslant 0, \\ (0.125,0,0.25)^{\mathrm{T}} & \text{if } x > 0. \end{cases}$$

The pressure P is given by $P = (\gamma - 1)(E - \frac{1}{2}\rho u^2)$ and $\gamma = 1.4$ is the ratio of specific heats of an ideal gas.

With this initial condition, the left domain and the right domain are separated by a contact discontinuity for $t > 0$. As time progresses, a rarefaction wave develops that propagates left on the left side of the contact discontinuity and the original shock propagates to the right on the right side of the contact discontinuity. This results in 4 discontinuities of the density profile as shown by the solid line in Fig. 8 for a particular time $t = 0.35$ [6,25].

We assume that we are given N Chebyshev coefficients b_k for the density profile shown in Fig. 8. We have used 600 Gauss quadrature points to evaluate these b_k coefficients from this profile. Thus one should note that these b_k are not exact. In Fig. 8, this Chebyshev approximation of $\rho(x)$ $(N = 23)$ is represented by the dotted line, which yields the Gibbs oscillations due to the multiple discontinuities. For the reconstruction each element of the matrix $\mathbf{T}$ was calculated using 600 Gauss quadrature points. In each of the five intervals, we use $m_s = 3, 7, 3, 3$, and 3, from left to right and thus $m = N = 23$. The reconstructed function obtained in this way is shown as the symbols in Fig. 7 and is within four significant figures of the original profile. This demonstrates the utility of the inverse method applied to multiple intervals.

7. Summary

This paper has presented a novel methodology for the resolution of the Gibbs phenomena that occurs when one expresses a piece-wise continuous, nonperiodic function as a Fourier series. The objective of any resolution is to recover the original function from the information contained in a finite set of Fourier coefficients. We have shown that the present approach, referred to as the inverse method, gives an exact resolution of the Gibbs phenomenon for polynomials. We have applied the inverse method to several polynomial and nonpolynomial test functions either on a single interval or

on multiple intervals. In numerous examples, a resolution to machine accuracy can be obtained with a small number of basis functions. An alternate method for the resolution of the Gibbs phenomenon was developed by Gottlieb and co-workers [14]. Their method, referred to as the direct method, is based on the Gegenbauer polynomials orthogonal with respect to weight function $(1-x^2)^{\lambda-1/2}$. In the comparisons of the inverse and direct methods that are presented in this paper, we have also provided a qualitative interpretation of the role of the parameter λ in the direct method.

Acknowledgements

This research is supported by a grant to BDS from the Natural Sciences and Engineering Research Council of Canada. BDS is also grateful for support from the French Embassy in Ottawa for a two month visit to the Department of Mathematics, University of Nice where this work was initiated. We also thank Dr. Richard Pasquetti and Dr. Robert F. Snider for their reading of the manuscript prior to publication and for their many constructive comments.

Appendix A. Analytic calculation of the $T_{k,\ell}$ and $V_{k,\ell}$ matrix elements

With the definition of T_{kl}, Eq. (16), we have that,

$$T_{k,l+1} - T_{k,l-1} = \int_{-1}^{1} \sin(k\pi x)[C_{l+1}^{\lambda}(x) - C_{l-1}^{\lambda}(x)]\,\mathrm{d}x, \tag{A.1}$$

where $k \neq 0$ and l is odd. With an integration by parts and the subsequent use of the differentiation formula,

$$\frac{\mathrm{d}}{\mathrm{d}x}[C_{l+1}^{\lambda}(x) - C_{l-1}^{\lambda}(x)] = 2(l+\lambda)C_{l}^{\lambda}(x) \tag{A.2}$$

we get that

$$T_{k,l+1} - T_{k,l-1} = (-1)^{k+1}\frac{2}{k\pi}[C_{l+1}^{\lambda}(1) - C_{l-1}^{\lambda}(1)] + \frac{2(l+\lambda)}{k\pi}V_{k,l}, \tag{A.3}$$

where

$$C_{l}^{\lambda}(1) = \frac{\Gamma(l+2\lambda)}{n!\Gamma(2\lambda)}. \tag{A.4}$$

Similarly, with the definition of V_{kl}, Eq. (21), for $k \neq 0$ and even l,

$$V_{k,l} - V_{k,l-2} = \int_{-1}^{1} \cos(k\pi x)[C_{l}^{\lambda}(x) - C_{l-2}^{\lambda}(x)]\,\mathrm{d}x \tag{A.5}$$

followed by an integration by parts and the use of Eq. (A.2), we get

$$V_{k,l} - V_{k,l-2} = -\frac{2(l-1+\lambda)}{k\pi}T_{k,l-1}. \tag{A.6}$$

With the specific values,

$$T_{k,l} = \begin{cases} 0 & \text{for } k = 0, \\ (-1)^{k+1} 4\lambda/(k\pi) & \text{for } k \neq 0,\ l = 1, \end{cases}$$

$$V_{k,l} = \begin{cases} 1 & \text{for } k = 0,\ l = 0, \\ 0 & \text{for } k \neq 0,\ l = 0, \\ \dfrac{1}{l+\lambda}\left[C_{l+1}^{\lambda}(1) - C_{l-1}^{\lambda}(1)\right] & \text{for } k = 0,\ l \neq 0, \end{cases} \tag{A.7}$$

the recurrence relations, Eqs. (A.3) and (A.6), can then be used to accurately calculate the matrix elements.

References

[1] R.A. Adomaitis, Spectral filtering for improved performance of discretization methods, Comput. Chem. Eng. 25 (2001) 1621–1632.
[2] R. Archibald, A. Gelb, Reducing the effects of noise in image reconstruction, J. Sci. Comput. 17 (2002) 167–180.
[3] C. Canuto, M.Y. Hussaini, A. Quarteroni, T.A. Zang, Spectral Methods in Fluid Dynamics, Springer Series in Computational Physics, Springer, New York, 1988.
[4] M. Cartwright, Fourier Methods for Mathematicians, Scientists and Engineers, Ellis Harwood, New York, 1990, pp. 67–69.
[5] H. Chen, B.D. Shizgal, A spectral solution of the Sturm-Liouville equation: comparison of classical and nonclassical basis sets, J. Comput. Appl. Math. 136 (2001) 17–35.
[6] R. Courant, K.O. Friedrichs, Supersonic Flow and Shock Waves, Springer, New York, 1976.
[7] P.J. Davis, P. Rabinowitz, Methods of Numerical Integration, Academic Press, New York, 1989.
[8] T.A. Driscoll, B. Fornberg, A Padé-based algorithm for overcoming the Gibbs phenomenon, Numer. Algorithms 26 (2001) 77–92.
[9] B. Fornberg, A Practical Guide to Pseudospectral Methods, Cambridge University Press, Cambridge, 1996.
[10] W. Gautschi, J. Comput. Appl. Math. 12 (1985) 61.
[11] A. Gelb, E. Tadmor, Detection of edges in spectral data II. Nonlinear enhancement, SIAM J. Numer. Anal. 38 (2000) 1389–1408.
[12] D. Gottlieb, J.S. Hesthaven, Spectral methods for hyperbolic problems, J. Comput. Appl. Math. 128 (2001) 83–131.
[13] D. Gottlieb, S.A. Orszag, Numerical Analysis of Spectral Methods: Theory and Applications, SIAM, Philadelphia, 1977.
[14] D. Gottlieb, C-W. Shu, On the Gibbs phenomenon and its resolution, SIAM Rev. 39 (1997) 644–668.
[15] D. Gottlieb, C-W. Shu, A. Solomonoff, H. Vandeven, On the Gibbs phenomenon 1: recovering exponential accuracy from the Fourier partial sum of a nonperiodic analytic function, J. Comput. Appl. Math. 43 (1992) 81–92.
[16] J.D. Jackson, Mathematics for Quantum Mechanics, Academic Press, New York, 1975, pp. 93–94.
[17] C. Lanczos, Discourse on Fourier series, Hafner Publishing Company, New York, 1966.
[18] R. March, P. Barone, Reconstruction of a piecewise constant function from noisy Fourier coefficients by Padé method, SIAM J. Appl. Math. 60 (2000) 1137–1156.
[19] R. Peyret, Spectral Methods for Incompressible Viscous Flow, Springer, New York, 2002.
[20] B.D. Shizgal, H. Chen, The quadrature discretization method (QDM) in the solution of the Schrodinger equation with nonclassical basis functions, J. Chem. Phys. 104 (1996) 4137–4150.
[21] G.A. Sod, A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws, J. Comput. Phys. 43 (1978) 1–31.

[22] E. Tadmor, J. Tanner, Adaptive mollifiers for high resolution recovery of piecewise smooth data from its spectral information, Found. Comput. Math. 2 (2002) 155–189.
[23] L.N. Trefethen, Spectral Methods in Matlab, SIAM, Philadelphia, 2000.
[24] L. Vozovoi, A. Weill, M. Israeli, Spectrally accurate solution of nonperiodic differential equations by the Fourier–Gegenbauer methods, SIAM J. Numer. Anal. 34 (1997) 1451–1471.
[25] G.B. Whitham, Linear and Nonlinear Waves, Wiley-Interscience, New York, 1974.